

**Supplemental Information**

**Genome-wide *In Vivo* CNS Screening**

**Identifies Genes that Modify CNS**

**Neuronal Survival and mHTT Toxicity**

**Mary H. Wertz, Mollie R. Mitchem, S. Sebastian Pineda, Lea J. Hachigian, Hyeseung Lee, Vanessa Lau, Alex Powers, Ruth Kulicke, Gurrein K. Madan, Medina Colic, Martine Therrien, Amanda Vernon, Victoria F. Beja-Glasser, Mudra Hegde, Fan Gao, Manolis Kellis, Traver Hart, John G. Doench, and Myriam Heiman**

## Supplemental Information

**Figure S1.** Characterization of Lentiviral Spread and Estimation of *in vivo* Lentivirus Multiplicity of Infection, Related to Figure 1.

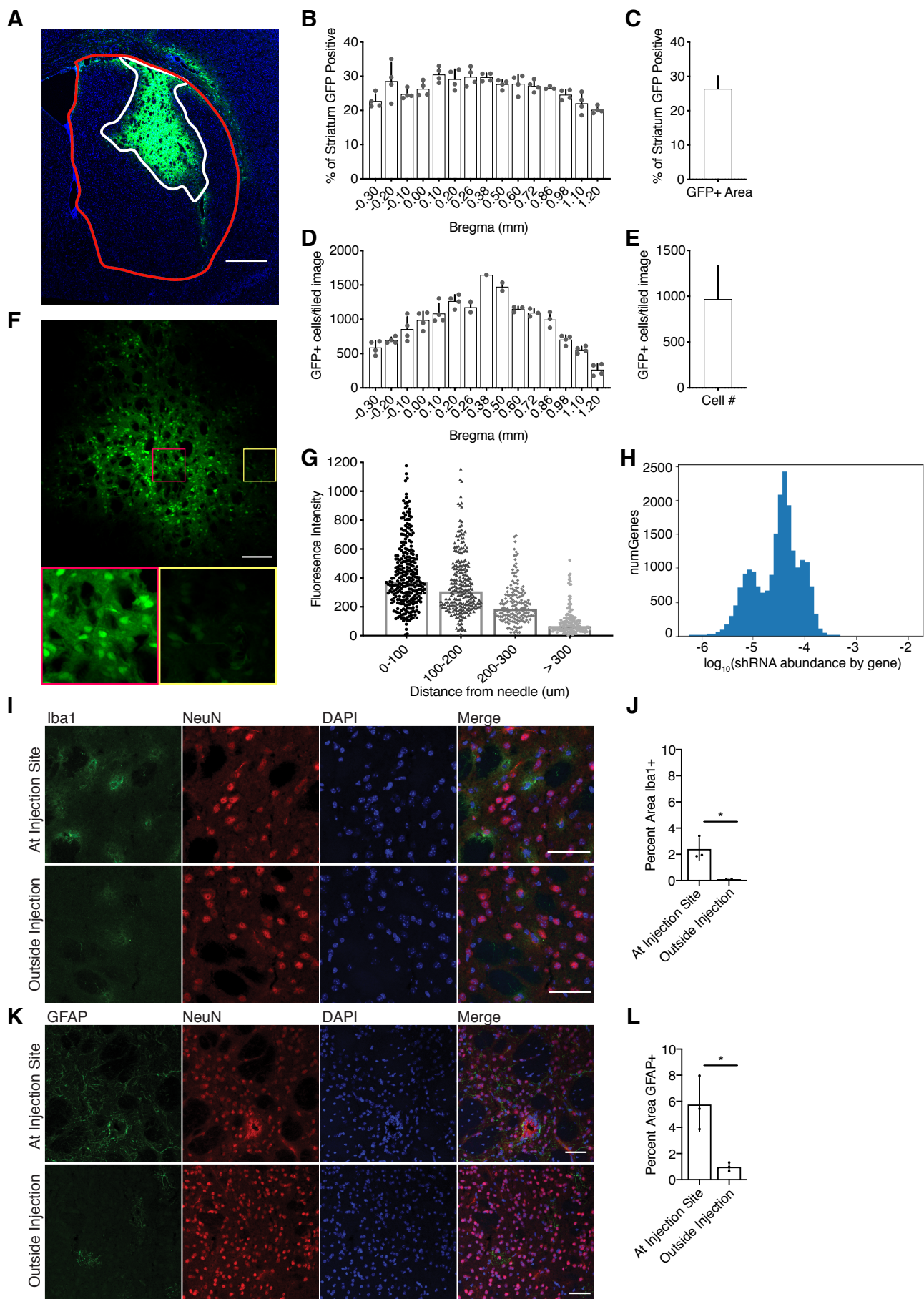
**Figure S2.** Coverage and Performance of shRNA and sgRNA Pooled Libraries *in vivo*, Related to Figure 1.

**Figure S3.** Comparison of Striatal-Enriched Genes and Core Essential Genes (CEG2) to Candidate Neuronal Essential Genes Identified in shRNA and CRISPR Screens, Related to Figures 2 and 3.

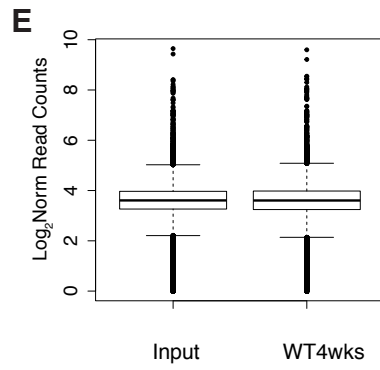
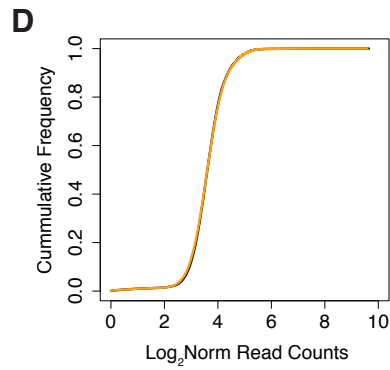
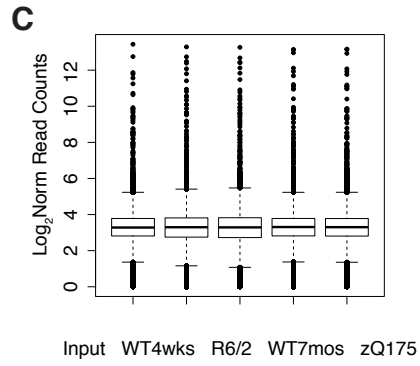
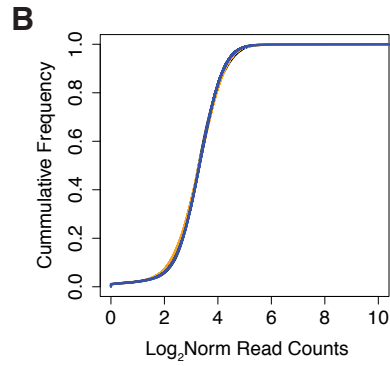
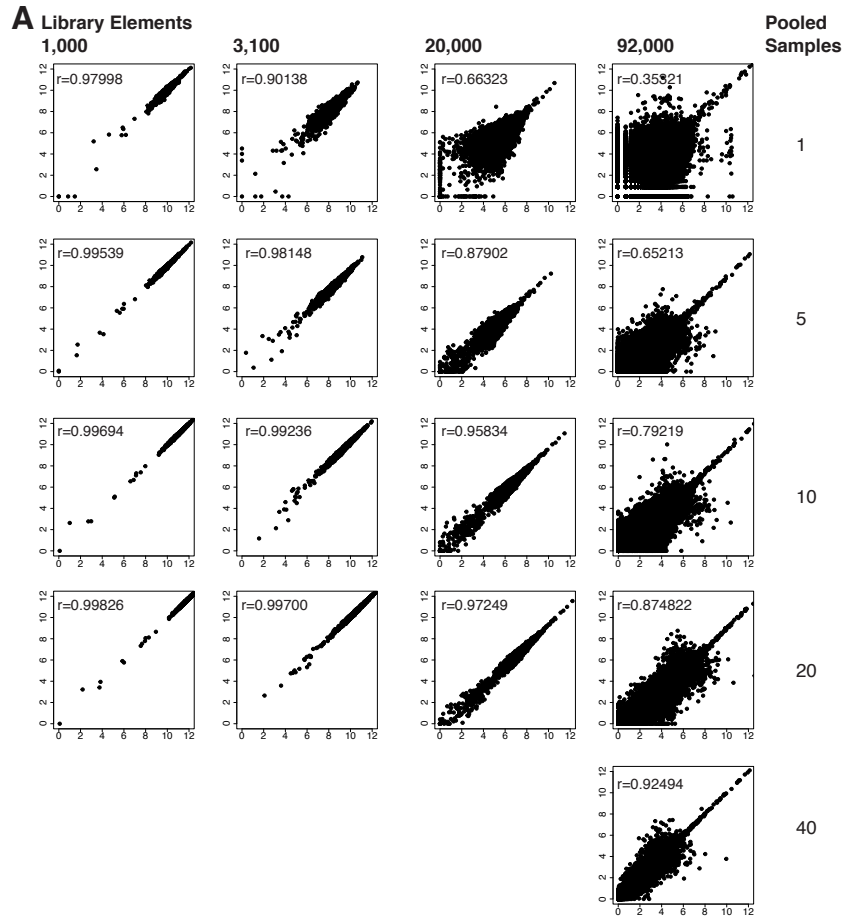
**Figure S4.** Identification of Candidate Protective and Vulnerability Factors in HD Mouse Models, Related to Figure 4.

**Figure S5.** Levels of *Nme1* Expression in Mouse and Human Brain, and Effects of *Nme1* knockdown and overexpression *in vivo* and *in vitro*, Related to Figure 5.

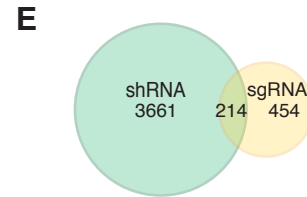
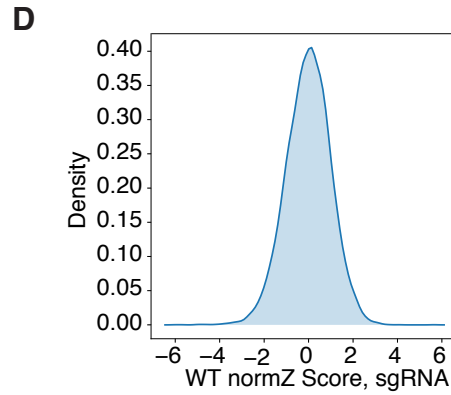
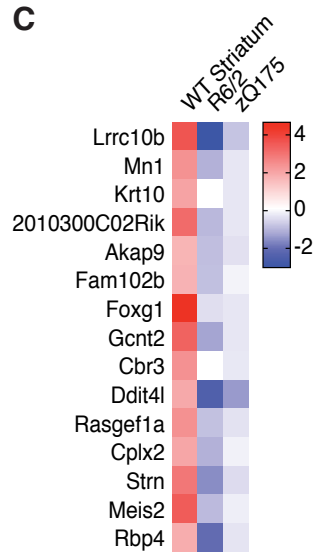
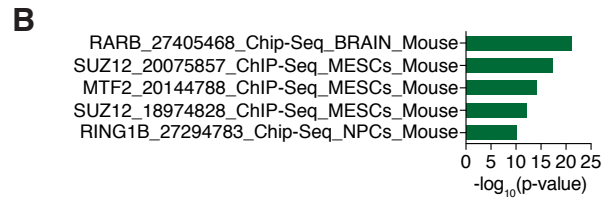
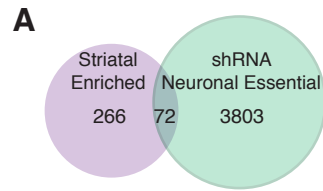




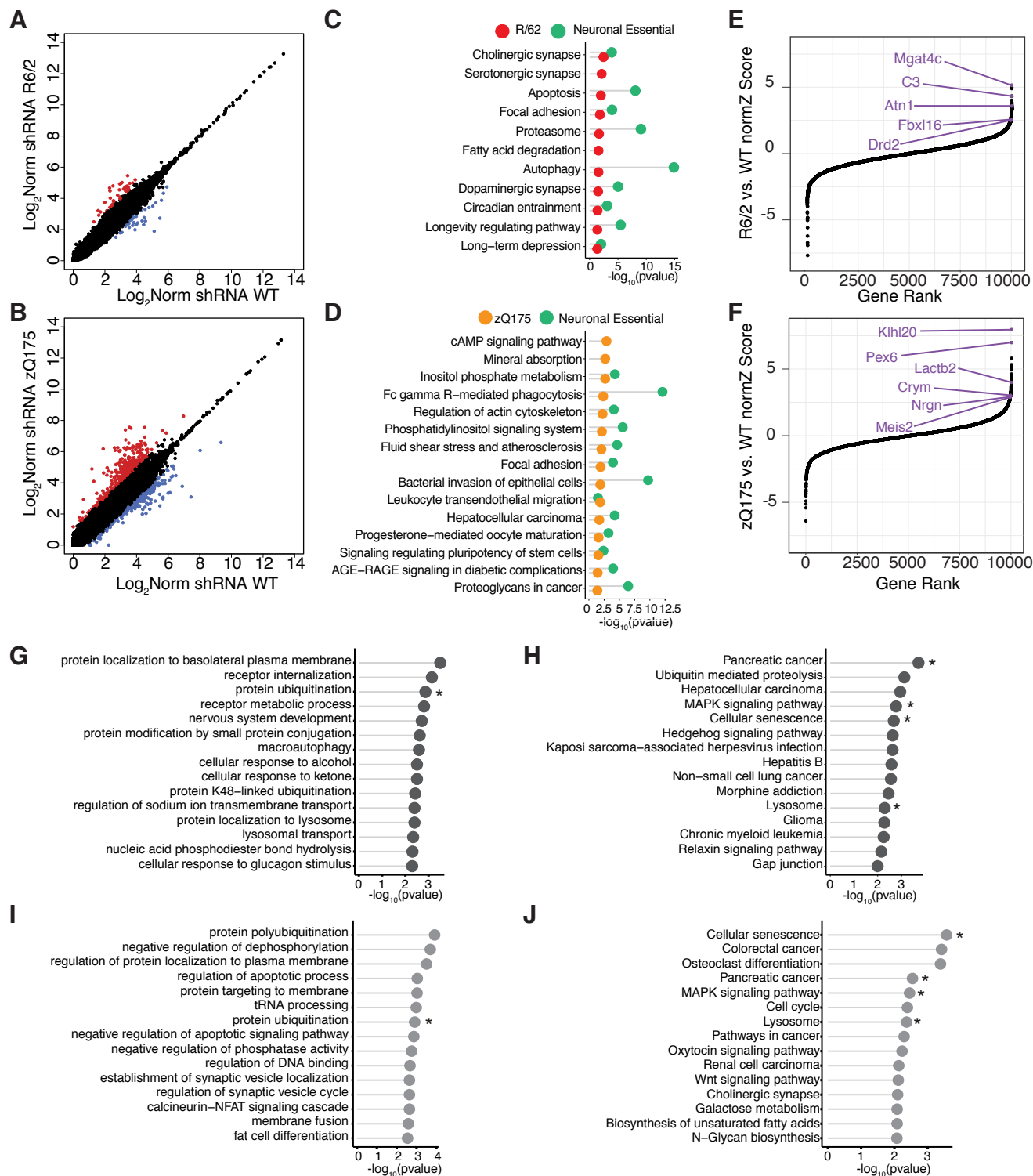
**Figure S1. Characterization of Lentiviral Spread and Estimation of *in vivo* Lentivirus Multiplicity of Infection, Related to Figure 1.** (A) Representative image of control, GFP cDNA lentiviral spread after injection into the adult striatum. Scale bar = 400 $\mu$ m. (B) Area of the striatum (red line in panel A) positive for GFP viral infection (white line in panel A) was measured for each section from anterior to posterior across the injection sites. The percent GFP-positive area present in 4 serial coronal sections is represented as mean and standard deviation (SD) for each coordinate relative to Bregma (mm). (C) The GFP positive area (effective viral coverage) across the injection sites is 26.41% of the striatum,  $n = 60$  sections. (D) GFP positive cells were manually counted in maximum intensity projections of serial 20x 10 $\mu$ m Z-stacks of the GFP positive injection area. The number of GFP positive cells for 4 serial sections is represented as mean + SD for each coordinate relative to Bregma (mm). (E) Bar graph of the average number of GFP-positive cells per 10 $\mu$ m Z-stack of a 20 $\mu$ m section,  $n = 56$  sections. Together the viral spread and number of cells counted per 10 $\mu$ m section indicate ~120,000 neurons are infected per striatal hemisphere (see STAR Methods for details). Bars represent mean + SD. (F) Representative image of endogenous fluorescence intensity of GFP-positive cells in a tissue section directly adjacent to the injection site (red box) vs. at the periphery of the injection (yellow box). Scale bar = 50 $\mu$ m. (G) Endogenous fluorescence intensity of individual cells was measured from summed intensity projections of tiled 20 $\mu$ m Z-stacks to evaluate *in vivo* MOI. Median fluorescence intensity in arbitrary units at the periphery is 65.32, while at the needle track is 371.00, indicating a 5.68-fold range in the distribution of viral integration events.  $n > 165$  cells per distance bin from 4 coronal sections, bars represent median value. (H) Histogram of the distribution of  $\log_{10}$  shRNA read counts by gene in the input library shows the relative abundance of the total shRNA targeting each gene. (I) Iba1, and (K) GFAP co-stained with NeuN and DAPI in virus-injected striatum of C57Bl/6J control mice. 40x z-stack images were taken at the center of the injection site and compared to a region peripheral to the injection site. (J/L) Quantification of Iba1 and GFAP staining showed a small localized increase in inflammatory markers adjacent to the center of the injection site, but not peripheral to the injection,  $n = 3$ , \*  $p$ -value  $< 0.05$ ; two-tailed student's  $t$ -test. Images have been modified to increase contrast for visualization purposes. All images were modified identically and quantification was performed on unmodified images as described in the methods.



**Figure S2. Coverage and Performance of shRNA and sgRNA Pooled Libraries *in vivo*, Related to Figure 1.** (A) Representative scatterplots of the log<sub>2</sub> normalized read counts in libraries of low to high complexity (1,000; 3,100; 20,000; and 92,000 elements). Pooling higher numbers of samples (1, 5, 10, 20, and 37) into a single replicate reduces variation and increases the Pearson correlation coefficient (r) as shown for each comparison. A single replicate is sufficient for full coverage and high correlation ( $r > 0.75$ ) of small 1,000 and 3,100 element libraries, while pooling of many samples per replicate is required for the 20,000 and 92,000 element libraries. (B) Cumulative density function (CDF) plot of the cumulative frequency vs. the log<sub>2</sub> normalized representation of shRNAs in the genome-wide library. The WT control replicates both 4 weeks (blue) and 7 months (orange) of *in vivo* incubation are similar to the plasmid input (black). (C) Boxplot of the distributions of log<sub>2</sub> normalized read counts in the shRNA screening data sets as compared to the plasmid input. (D) CDF plot of the cumulative frequency vs. the log<sub>2</sub> normalized representation of sgRNAs in the input library (black) and the WT 4 weeks samples (orange). (E) Boxplot of the distributions of log<sub>2</sub> normalized read counts in the sgRNA screening data sets as compared to the plasmid input. Together these plots illustrate recovery of the library at a depth sufficient for analysis of differential library representation in the pooled replicate samples.

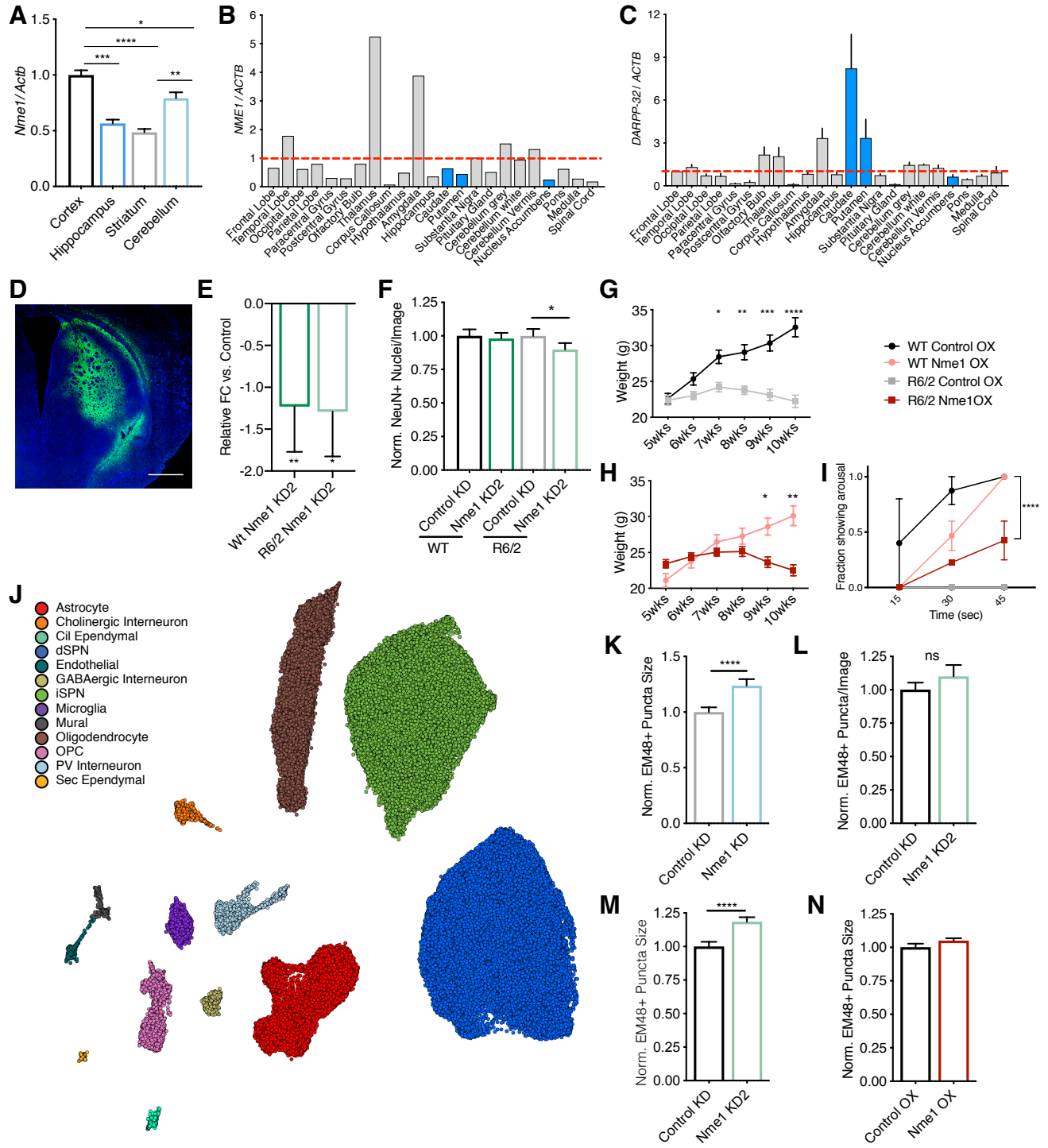


**Figure S3. Comparison of Striatal-Enriched Genes and Core Essential Genes (CEG2) to Candidate Neuronal Essential Genes Identified in shRNA and CRISPR Screens, Related to Figures 2 and 3.** (A) Venn diagram of overlap between striatal enriched genes [ $>1.5 \log_2$  fold-change enrichment in striatum vs. rest of brain (Kasukawa et al., 2011)] and neuronal essential genes identified in the WT shRNA screen. (B) ChEA analysis of the 72 neuronal essential genes which display  $> 1.5 \log_2$  fold-change striatal enriched expression (overlap in panel A, and listed in Table S2), to identify putative regulators of these genes from published chromatin immunoprecipitation studies. Results represented as Fisher's exact test  $-\log_{10}$  p-value. (C) Heatmap of the top 15 candidate neuronal essential genes with enriched expression in the striatum vs. the rest of the brain (left column, fold enrichment in striatum). All of these genes have significantly decreased expression (corrected p-value  $< 0.05$ ) in the striatum of the either the R6/2 or zQ175 HD mouse models, or both (right two columns, fold downregulation in HD model). (D) Density plot of normZ scores for the WT CRISPR screen shows normal distribution. (E) Venn diagram of the overlap of neuronal essential genes identified by both the shRNA and CRISPR sgRNA screens, hypergeometric p-value = 0.023.



**Figure S4. Identification of Candidate Protective and Vulnerability Factors in HD Mouse Models, Related to Figure 4.** (A and B) Scatterplots (corresponding to Figure 4B and 4D) illustrating the correlation of  $\log_2$  normalized read counts in the R6/2 and zQ175 screens respectively. Points represent individual shRNA hairpins enriched (red) or depleted (blue) with  $> 2$ -fold differential shRNA sequencing representation in the mutant as compared to the isogenic WT control. Highly correlated sequences with a  $\log_2$  normalized value greater than 8 are shRNAs that are highly represented in the library, but are not differentially recovered in the screen. (C and D) Top KEGG pathways significantly associated with candidate essential genes (knockdown is harmful, library elements are depleted as described in the main text) unique to either the R6/2 (red) and zQ175 (orange) screens respectively, and not identified as essential in the WT neuronal essential shRNA screen, represented with Fisher's exact test  $-\log_{10}$  p-value. Even though these genes were not identified as essential in the WT neuronal essential shRNA screen, there was still pathway overlap evident with WT neuronal essential genes (green). (E and F) Plot of normZ scores vs. rank of candidate mHTT vulnerability factors (knockdown is protective, library elements in this case are enriched) identified by DrugZ. The top genes from each arm of the screen are labeled in purple. Top (G and I) KEGG pathways and (H and J) GO terms significantly associated with candidate vulnerability factors in the R6/2 and zQ175 screens respectively, Fisher's exact test expressed as  $-\log_{10}$  p-value. Pathways overlapping between the R6/2 and zQ175 screens are marked with an \*.





**Figure S5. Levels of *Nme1* expression in mouse and human brain, and effects of *Nme1* knockdown and overexpression *in vivo* and *in vitro*, Related to Figure 5.** (A) qRT-PCR of *Nme1* mRNA expression in WT mouse brain regions normalized to *Actb* shows lowest expression of *Nme1* in striatum,  $n = 3$  biological replicates. Ordinary one-way ANOVA \*\*\*\*  $P < 0.0001$ , Tukey's multiple comparison test \*  $p = 0.027$ , \*\*  $p = 0.003$ , \*\*\*  $p = 0.0003$ , \*\*\*\*  $p < 0.0001$ . (B) Multiplex Taqman qRT-PCR analysis of a human cDNA array (Origene) confirms low expression of *NME1* (normalized to *ACTB*) in the caudate, putamen, and nucleus accumbens (blue) as compared to overall human brain expression;  $n = 1$  array. (C) Enriched expression of *DARPP-32* (*PPP1R1B*) in the caudate and putamen (normalized to *ACTB*) in the human cDNA array confirms the specificity of the array panel;  $n = 3$  arrays. (D) 10x tiled image of a representative GFP AAV9 viral spread area in the striatum upon stereotaxic intracranial injection (Scale bar = 300 $\mu$ m). Contrast in GFP and DAPI images was increased for visualization purposes. (E) qPCR quantification of the relative fold-change of *Nme1* mRNA expression in the *Nme1* KD2 as compared to Control KD injected hemispheres to confirm knockdown with a second *Nme1* shRNA hairpin sequence,  $n = 10$ , \*\* =  $p$ -value = 0.0097, \*  $p$ -value = 0.013; one-tailed, paired student's  $t$ -test of DeltaCt values. (F) Quantification of NeuN-positive cells in the Control KD vs. *Nme1* KD2 injected striatal hemispheres of WT ( $n = 4$ ) and R6/2 ( $n=3$ ) mice. Points are an average of three images per section from 8 or 16 sections; one-tailed, paired  $t$ -test,  $p$ -value = 0.01. (G) Effect of striatal Control OX on the body weights of WT or R6/2 HD model mice. Two-way repeated measures ANOVA  $p$ -value genotype  $<0.0001$  with Sidak's, multiple comparison test, \*  $p = 0.013$ , \*\*  $p = 0.0046$ , \*\*\*  $p = 0.0006$ , \*\*\*\*  $P < 0.00001$ . (H) Effect of striatal *Nme1* OX on body weight of WT vs. R6/2 HD model mice. Two-way repeated measures ANOVA with Sidak's multiple comparison test, \*\*  $p = 0.0013$ , and \*\*\*\*  $P < 0.0001$ . (I) Effect of striatal Control OX or striatal *Nme1* OX on the arousability (at 9 weeks of age, 4 weeks after viral transduction) of WT or R6/2 HD model mice. Two-way repeated measures ANOVA  $p$ -value genotype = 0.017 with Sidak's, multiple comparison test, \*\*\*\*  $p$ -value  $<0.0001$  WT Control OX and WT *Nme1* OX vs. R6/2 Control OX. (J) ACTIONet graphs of the major annotated cell types in the R6/2 model of HD ( $n = 95,885$  nuclei across 20 mice: 5 WT and 5 R6/2 model mice, each for Control and *Nme1* OX). (K) Quantification of EM48+ puncta size in Control KD vs. *Nme1* KD injected hemispheres of R6/2 mice ( $n = 3$ ). (L and M) Quantification of mHTT aggregates as normalized EM48+ puncta per image and size of EM48+ puncta in Control KD vs. *Nme1* KD2 injected hemispheres of R6/2

mice ( $n = 3$ ). Panels K-M, points are an average of three images per section from 16 sections, \*\*\*\* p-value  $< 0.0001$ ; one-tailed paired student's  $t$ -test. **(N)** Quantification of EM48+ puncta size in Control OX vs. Nme1 OX injected hemispheres of R6/2 mice ( $n = 4$ ). Points are an average of three images per section from 8 sections; n.s. one-tailed paired student's  $t$ -test. All error bars represent mean and  $\pm$  SEM.